

テキストによるコミュニケーションにおける影響の普及モデル

Influence Diffusion Model in Text-Based Communication

松村 真宏

Naohiro Matsumura

東京大学大学院工学系研究科 / 科学技術振興事業団

School of Engineering, The University of Tokyo / Japan Science and Technology Corporation
matumura@miv.t.u-tokyo.ac.jp, <http://www.miv.t.u-tokyo.ac.jp/~matumura/>

大澤 幸生

Yukio Ohsawa

筑波大学大学院ビジネス科学研究科 / 科学技術振興事業団

GSSM, University of Tsukuba / Japan Science and Technology Corporation
osawa2@gssm.otsuka.tsukuba.ac.jp, <http://www.gssm.otsuka.tsukuba.ac.jp/staff/osawa/>

石塚 満

Mitsuru Ishizuka

東京大学大学院情報理工学系研究科

School of Information Science and Technology, The University of Tokyo
ishizuka@miv.t.u-tokyo.ac.jp, <http://www.miv.t.u-tokyo.ac.jp/~ishizuka/>

keywords: influence diffusion model, text-based communication, BBS

Summary

Business people, especially marketing researchers, are keen to understand peoples' potential sense of value to create fascinating topics stimulating peoples' interest. In this paper, we aim at finding influential people, comments, and terms contributing the discovery of such topics. For this purpose, we propose an Influence Diffusion Model in text-based communication, where the influence of people, comments, and terms are defined as the degree of text-based relevance of messages. We apply this model to Bulletin Board Service(BBS) on the Internet, and present our discoveries on experimental evaluations.

1. はじめに

近年、インターネット上の電子掲示板からユーザの潜在的なニーズや価値観を発掘して、魅力的な商品の開発や画期的なマーケティング戦略に役立たせることが注目を浴びてきている [博報堂 00]。これは、電子掲示板にコメントを投稿している人はそこでの話題に強い関心のある人が多く、その結果としてユーザの潜在的なニーズや価値観が議論されていることが多いからである。また、ユーザ同士の議論が商品に新たな価値を生み出すことや [国領 00]、購買意思決定に大きな影響を与えることも指摘されている [池田 97]。

そこで本論文では、ある人の発した語への興味が他の人に伝播していくプロセスに着目し、テキストによるコミュニケーションにおける影響の普及を表すモデルを提案する。このモデルでは、ある話題がコミュニティに投げ込んだ文脈に、コミュニティ中で他のコメントたちの内容が強く支配されていくときに、この話題を盛り上げる話題、そのような話題の提起者をオピニオンリーダー [Katz 55, Rogers 62] と見なす。このように文脈の支配の強さに焦点をあてて支配力の強いコメントを見出すのは、社会の現状において人々が「言われてみれば」興味を寄せる潜在的な話題やその話題を提起した人を把握する

のに役立つと考えられるからである。

以下では、2章で影響の普及モデルの詳細、3章で実験条件について述べた後、4章、5章で電子掲示板に適用した例を示し、6章で実験結果の議論を行う。

2. 影響の普及

2.1 コミュニケーションにおける影響の普及

電子掲示板ではテキストを媒介としたコミュニケーションが行われているので、そこでやり取りされる情報は必然的に文字で表現される。したがって、コミュニケーションにおいて文字すなわち語への興味が伝播していく過程を観察すれば、他のコメントの内容を強く支配するような影響力のあるコメント、語、オピニオンリーダーを見つけることができよう。

電子掲示板上のコメントは誰でも読むことができるので、厳密に誰がどのコメントから影響を受けたのかを知ることにはできない。しかし、ある人 x が投稿したコメント C_x に返信している人 y は、 C_x を読んで何らかの影響を受けた人である可能性が高い。また、 C_x の影響は、伝播先のコメントを介して更にその先にまで、 C_x を起点とするコメントチェーン（コメントの返信関係の連なり）上を連鎖的に伝播することになり、このコメントチェーンが長く続けば C_x の影響は広範囲に及ぶと考えられる。

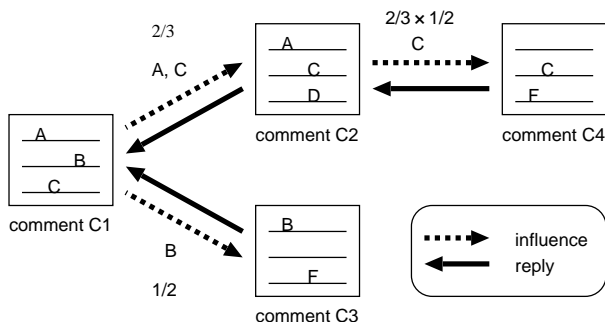


図 1 C_1 が他のコメントに及ぼす影響量 .

ここで、コメントチェーン上を伝播していく語の割合に着目すると、 C_x が他のコメントに及ぼす影響量を定式化することができる。まず、コメント C_x 中の語の集合を w_x 、 C_x に返信しているコメント C_y 中の語の集合を w_y とすると、 C_x から C_y に伝播した影響量 $i_{x,y}$ は次のようになる。

$$i_{x,y} = \frac{|w_x \cap w_y|}{|w_y|} \tag{1}$$

また、 C_y に C_z がさらに返信している場合において、 C_x から C_z に伝播した影響量 $i_{x,z}$ は次のようになる。

$$i_{x,z} = \frac{|w_x \cap w_y \cap w_z|}{|w_z|} \cdot i_{x,y} \tag{2}$$

式 (1) は文脈支配の関係を継承する単語の比率で表しており、式 (2) はこの比率をコメントチェーンに沿って掛け合わせていくことを表している。これは、ある単語がある人にとって支配的な文脈になり、その支配的な文脈がさらにその次にメッセージの伝搬する人にとって支配的になっていく比率を求めることを意味している。以上の定式化により、 C_x が他のコメントに及ぼす影響量を測ることができるようになる。

例えば、図 1 で表されるコメント C_1, C_2, C_3, C_4 からなるコメントチェーンを考える。 C_1 から C_2 には語 A, C が、 C_1 から C_3 には語 B が伝播しており、さらに C_2 から C_4 へ語 C が伝播している。このコメントチェーンにおいて、 C_1 が他の各コメントに与える影響量は次のようになる。

C_1 から C_2 へ伝わった影響量： C_1 から C_2 へ伝わった語は 2 語であり、 C_2 が発している語はこれを含む 3 語であるから、 C_1 から C_2 へ伝わった影響量は $2/3$ 。

C_1 から C_3 へ伝わった影響量： C_1 から C_3 へ伝わった語は 1 語であり、 C_3 が発している語はこれを含む 2 語であるから、 C_1 から C_3 へ伝わった影響量は $1/2$ 。

C_1 から C_4 へ伝わった影響量： C_2 から C_4 へ伝わった語は C_2 が C_1 から受け取った 2 語のうち 1 語であり、また C_1 から C_2 へ伝わった影響量は $2/3$ であるから、 C_1 から C_4 へ伝わった影響量は $2/3 \cdot 1/2 = 1/3$ 。

ここで、あるコメントの発した影響量を、そのコメントが他のコメントに及ぼした影響量の総和になると定義する。すると、図 1 のコメントチェーンにおいて C_1 が発した影響量は (C_1 から C_2 へ伝わった影響量) + (C_1 から C_3 へ伝わった影響量) + (C_1 から C_4 へ伝わった影響量) = $2/3 + 1/2 + 1/3 = 3/2$ となる。同様の手続きにより、 C_2 が発した影響量は $1/2$ 、 C_3 が発した影響量は 0、 C_4 が発した影響量は 0 と計算される。したがって、図 1 のコメントチェーンにおいて最も影響力をもつコメントは C_1 となる。

本研究で対象とする電子掲示板では、影響を伝播する媒体は、コメント、人、語の 3 種類ある。そこで、各々の媒体が他の媒体に及ぼす影響力を媒介影響量と定義し、以下でその詳細について述べる。

2.2 コメントの媒介影響量

まず、コメント C_i の媒介影響量を求める。2.1 節での議論から、コメント C_i からコメント C_z に至るコメントチェーン $\xi_{i,z} = \{C_i, C_j, C_k \dots C_p, C_q, C_r \dots C_x, C_y, C_z\}$ $\{i < j < k \dots p < q < r \dots x < y < z\}$ において、 C_i が C_r に及ぼした影響量 $i_{i,r}$ は式 (1)(2) を拡張した式 (3) で表される。

$$i_{i,r} = \frac{|w_i \cap w_j \cap \dots \cap w_r|}{|w_r|} \cdot i_{i,q} \tag{3}$$

ただし、 w_r は C_r 中の語の集合であり、 $\{w_i \cap w_j \cap \dots \cap w_r\}$ は C_i から C_r まで伝播した語の集合である。式 (3) は、 C_r が発した語のうち C_i から受け取った語の割合だけ $i_{i,q}$ の影響量が $i_{i,r}$ に伝播することを意味している。すると、このコメントチェーン $\xi_{i,z}$ 上を伝わる C_i の影響量 $I_{\xi_{i,z}}$ は、

$$I_{\xi_{i,z}} = i_{i,j} + i_{i,k} + \dots + i_{i,y} + i_{i,z} \tag{4}$$

となる。ここで、 C_i を起点と見なしたときのコメントチェーン (例えば ξ と書く) の全ての集合 τ_i についての I_{ξ} の総和が C_i が発した媒介影響量 D_{C_i} となるから、

$$D_{C_i} = \sum_{\xi \in \tau_i} I_{\xi} \tag{5}$$

となる。

2.3 人の媒介影響量

次に、投稿者 x の媒介影響量 D_x は x が発言したコメントの媒介影響量の総和となるから、

$$D_x = \sum_{i \in \kappa_x} D_{C_i} \tag{6}$$

となる。ただし、 κ_x は x が発言した全てのコメントの集合である。



図 2 実験に用いた電子掲示板 .

2.4 語の媒介影響量

最後に、コメントチェーン上を伝播していく語の媒介影響量を求める。2.2 節における $i_{i,r}$ は、式 (3) より語の集合 $\{w_i \cap w_j \cap \dots \cap w_r\}$ を介している。ここで、各語が均等に影響を伝播していると考え、1 語あたりの影響量すなわち $t \in \{w_i \cap w_j \cap \dots \cap w_r\}$ なる語 t が C_i から C_r へ伝える影響量 $j_{i,r,t}$ は式 (7) で表される。

$$j_{i,r,t} = \frac{1}{|w_i \cap w_j \cap \dots \cap w_r|} \cdot i_{i,r} \quad (7)$$

すると、コメントチェーン $\xi_{i,z}$ を伝わる t の影響量 $J_{\xi,t}$ は

$$J_{\xi,t} = j_{i,j,t} + j_{i,k,t} + \dots + j_{i,y,t} + j_{i,z,t} \quad (8)$$

となる。ここで、語 t の媒介影響量 D_t は、 t を含む全てのコメントチェーン ξ_t についての $J_{\xi,t}$ の総和であるから、

$$D_t = \sum_{\xi \in \xi_t} J_{\xi,t} \quad (9)$$

となる。

3. 実験条件

3.1 電子掲示板

電子掲示板には、コンテンツを重視せずにユーザー同士の自由な交流を促進する娯楽タイプの掲示板から、ユーザーの立場から商品・サービスが評価されたり口コミの情報提供がされるユーザーによる商品・サービス評価タイプの掲示板まで様々なタイプがあるが、ユーザの潜在的なニーズや価値観を知るためには後者の商品・サービス評価タイプが望ましい [石川 00]。そこで本研究では、我々の生活に浸透しているユニクロに関する情報交換が

行われている図 2 の Yahoo! の掲示板*1 から、1999 年 12 月 3 日 ~ 2001 年 7 月 27 日までに 430 人により投稿された 1161 件のコメントを対象として分析を行った。

なお、コメント中の引用文（多くの場合「>」で始まる行）が多くなればそれだけ伝播してくる語が増えるので、媒介影響量が大きくなる懸念がある。そこで、予めこの電子掲示板における引用文の使われ方を調べてみると、相手の文章を引用する理由には、引用元のコメントの意見を支持する場合と、同じことを自分の言葉で書くかわりに相手のコメントを引用して自分の意見とする場合に分類できることが分かった。また、全文を引用して残す場合は 1161 件のうち 2 件だけしかなく、逆に引用しないでコメントを書く場合はもとの文章の単語は文脈に応じてそのまま用いることが多く、同じ意味の言葉を別の表現に置き換えることはほとんどなかった。したがって、もとのコメントを引用する場合もしない場合も、新しいコメントでは自分の文脈にとって必要なだけの単語はもとのコメントから引用していると解釈できるので、本論文ではコメント全文を解析に用いた。

なお、コメントは日本語で書かれているため、予め前処理として茶筌 [茶筌] による形態素解析により名詞、形容詞、副詞、動詞、未知語だけを残し、そこから「です」や「ます」など通常はキーワードになり得ないストップワードを除いた。

3.2 評価方法

実際のマーケティングなどビジネス目的で電子掲示板のコメントを解析する現場では、特別に訓練された分析者が掲示板中の全ての発言に目を通して重要な発言やオピニオンリーダーをピックアップしている。しかし、この判断基準は分析者の経験や感性に基づいているので、盛り上がった議論、オピニオンリーダーを統一的な定義で扱い統一的な指標で評価することは現状では困難である。また、議論が盛り上がることの客観的な指標として、従来からコメントへの返信数やコメントチェーンの大きさを見ることが行われているが、形式的には返信していてもそれまでの話題の流れを踏まえずに新しい内容を展開している場合も多く、それだけでは議論の盛り上がりは判断できない。

したがって、最終的には個々のコメントの内容に目を通した上で議論が盛り上がっているか否かを判断しなければ本当に議論が盛り上がっているのかを判断できない。そこで本研究では、著者らがコメントの内容を一つ一つ精読し、話題が後の盛り上がりにも寄与する度合いを判断することにより実験結果の評価を行った。そして、研究に参加していない研究者、社会心理学者、社会人学生、学生を含めて数人で実験結果を議論し、実験結果の評価の

*1 <http://messages.yahoo.co.jp/bbs?.mm=GN&action=1&board=1834999&tid=a5fa5ka5afa5m9a5a4ada4jbfmbd8a4dea4la1aa&sid=1834999&mid=1>

妥当性を確認した。

4. 電子掲示板の分析事例 1

本章ではまず 17 件のコメントからなるコメントチェーンを分析する。このコメントチェーンで話題になっているトピックは主にフリースであり、人気の色やネット通販などの話題で盛り上がっている。全 1161 件のコメントの分析については 5 章で後述する。

4.1 盛り上がる話題を提供するコメントの発見

2.2 節で提案したコメントの媒介影響量に基づいて得られた結果を表 1、図 3 に示す。

表 1 図 3 におけるコメントの媒介影響量の上位 5 つ

順位	コメント ID	媒介影響量
1	#604	1.504
2	#615	0.574
3	#618	0.375
4	#614	0.337
5	#605	0.237

表 1 のコメントを順番に見ていくと、媒介影響量第 1 位として選ばれたコメント #604 の要旨は

[コメント #604 の要旨] フリースの週末限定セールに行ってきました。限定色のピンクがサイズによってはやばそうでした。パールオリブが意外ときれいだったので買ってしまいました。フルジップの方が人気がありましたが、私はーフジップの方がかわいいと思いました。

となり、このコメントチェーンでの話題の多くを提供していた。ここで用いたコメントチェーンはコメント #604 を起点としているので、#604 が選ばれるこの結果は妥当であろう。

媒介影響量第 2 位のコメント #615 は、コメント #604 への返信であり、その要旨は次のようである。

[コメント #615 の要旨] 私もパールオリブが気になっているのですが、チラシの色とネットでの色がかなり違うように思うのですが、実物はどちらの色に近いでしょうか？

ここで述べられている色に関する話題も、このコメントチェーンの中核を成している話題であり、コメント #618、#619 も引き続きこの話題で盛り上がっていた。

媒介影響量第 3 位のコメント #618 もまたフリースの色に関するコメントである。以下の要旨から分かるように、コメント #618 はコメント #619 へ大きな影響を与えていた。

[コメント #618 の要旨] フリースのパールオリブ、私もネットで見てすごく気に入りました。でもチラシと色が違ったので「あれ？」と思ったんですが、実物はチラシと同じ色でした。やっぱり実物見ないとわからないですね。

[コメント #619 の要旨] なるほど。パールオリブの色はやっぱりチラシの色なんだ。
> やっぱり実物見ないとわからないですね。本当ですね。わたしも出来る限りそうしようと思います。

媒介影響量第 4 位のコメント #614 の要旨は

[コメント #614 の要旨] ユニクロ店員です。ピンクのフリースは一日で完売でした。実物を見たら安心して通販できる気がしました。

であり、図 3 の左半分のコメントチェーンの起点となる話題を提供していた。

媒介影響量第 5 位のコメント #605 はコメント #604 への返信であり、その要旨は

[コメント #605 の要旨] 脱ぎ着しやすいからフルジップの方が人気なのかな。今年はピーチとウルトラマリンを買ったのですが、両方ともきれいな色でお気に入りです。店員さんによると、ピンクは昼くらいには売り切れたそうです。

である。ここでも色に関する話題がコメント #606、#607 に引き継がれ盛り上がっていた。

従来のコメントチェーン解析においては、被返信数の多いコメント、コメントチェーンの大きさ、発言の大きさ（単語数の多さ）を話題の盛り上がりの尺度として用いられることが多い。図 3 のコメントチェーンでは、コメント #604、#614、#605、#615 の被返信数はそれぞれ 4、3、2、2、それ以外のコメントの被返信数は 1 であった。しかし、コメントが引き起こす話題の盛り上がりとその文脈への支配関係を見ると、被返信数 3 のコメント #614 よりも被返信数 2 のコメント #615 の方が影響力が大きかった。また、コメントチェーンを眺めるとコメント #618 よりもコメント #614 の方がそのあとのコメントチェーンが長い。コメント #614 への返信という形をとるコメント #629 や #631 はコメント #614 と別の文脈に話を展開しており、コメント #619 に大きな影響を残しているコメント #618 の方が文脈に対する影響力が大きかった。また、発言の最も大きいコメントは #607 であるが、#607 へは誰も返信しておらず他のコメントへの影響力はなかった。したがって、媒介影響量を用いた方が正確に盛り上がる話題を提供するコメントを取り出せていたことがわかる。

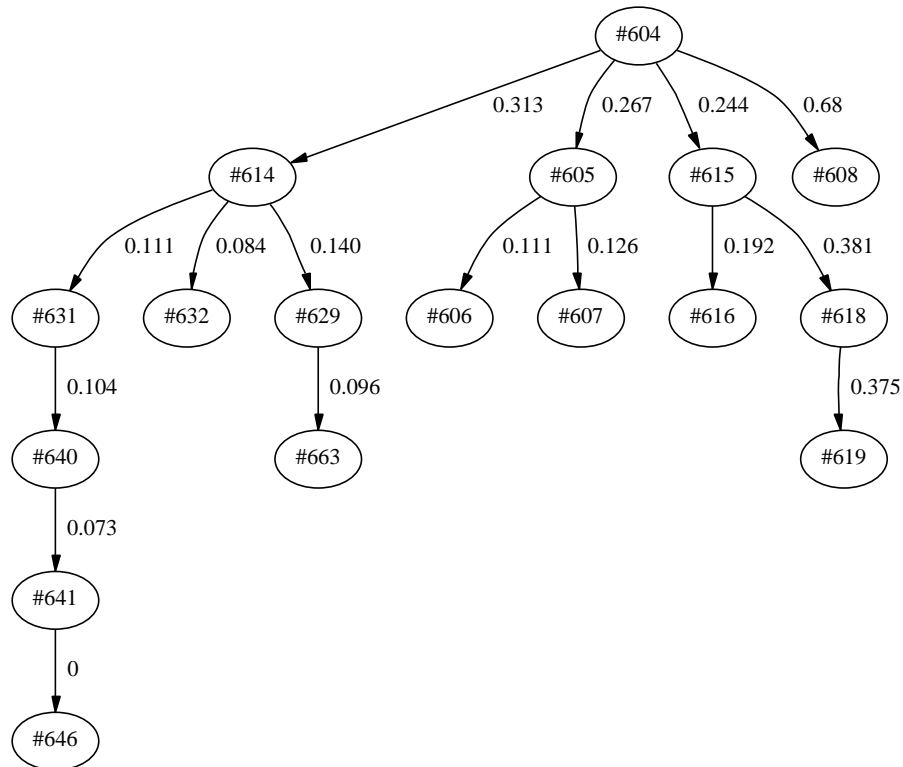


図 3 実験に用いたコメントチェーン．図中の有向リンクは影響の伝播する向き，ノード中の数字はコメントの ID 番号，リンク横の数字はコメントの媒介影響量を表している．

4.2 オピニオンリーダーの発見

2.3 節で提案した人の媒介影響量に基づいて得られた結果を図 4，表 2 に示す．表 2 を見ると，4.1 節で盛り上がる話題を提供していることを既に確認したコメント #604，#615，#618，#614，#605 の投稿者が漏れなく選ばれており，オピニオンリーダーを正確に取り出せると判断できる．

表 2 図 4 における人の媒介影響量の上位 5 名

順位	メンバー ID	(コメント ID)	媒介影響量
1	M073	(#604)	1.504
2	M049	(#615, #619)	0.574
3	M193	(#618)	0.375
4	M009	(#614, #663)	0.337
5	M010	(#605)	0.237

なお，被返信数やコメントチェーンの長さなどのコメントチェーンの構造に着目した方法もあるが，M193 は被返信数 1 でそのあとのコメントチェーンも短いコメント #618 の投稿者であるので，そのような手法で取り出すことは難しい．また，発言数の多い人をオピニオンリーダーとみなす考えもあるが，M086 は最も発言数が多い（3 回）にも関わらず，媒介影響量は低い．その理由を確かめるために，M086 が発言したコメントの内容を見てみると次のようであった．

[コメント #629, #631 の要旨] ピンクのフリー

スをネットで買いました．街中で着用するとジロジロ見られるんだよねえー．ところで，私の boyfriend がユニクロで働きたいと長年切望しているのですが，何でもいいので情報頂ければと...（注：コメント #629，#631 はほとんど同じ内容であったので一つにまとめた）

[コメント #641 の要旨] レスありがとうございます．まずはバイトか準社員からがいいかな...面接ではどんな事聞かれるんですか？絶対受かるレシピ!? とか...ねえよなあ...．

M086 の発言の内容も，このコメントチェーンの文脈に沿ったものであり，またユニクロで働くにはどうすればいいのかという新しい話題を提供していた．しかし，コメント #629，#631 でのピンクのフリースに関する話題はコメント #604，#605，#614 で既に十分な盛り上がりを見せていたために，それより後に M086 が発言した時には参加者の興味は他に移っていたと考えられる．また，ユニクロで働くことに関しては 2 件の返信があったが，その後は全く盛り上がりなかったことを考えると，この掲示板を見ている人の多くはユニクロで働くことよりもむしろユニクロの商品に興味があったために，盛り上がるような話題にならなかったのであろう．

4.3 盛り上がる話題を表す語の発見

次に，2.4 節で提案した語の媒介影響量を用いて，コメントチェーンから盛り上がる話題を表す語を取り出す

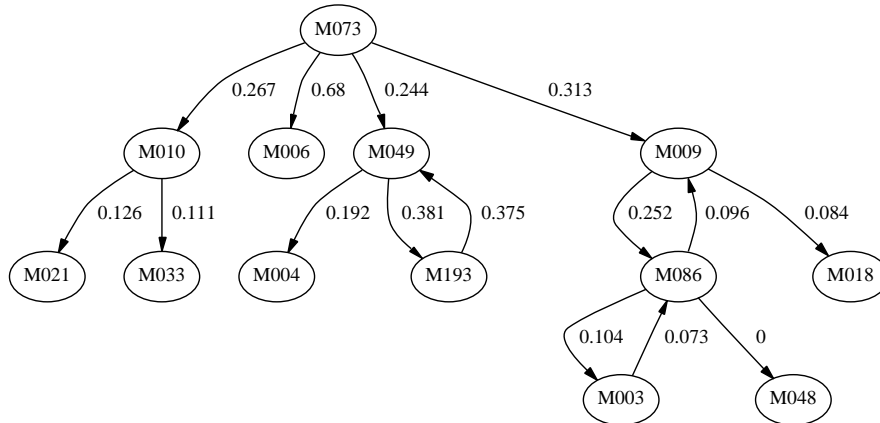


図 4 図 3 のコメントの投稿者間の返信関係を表すヒューマンネットワーク．ノード中のラベルは発言者の ID 番号，リンク横の数字は人の媒介影響量を表している．

ことを試みる．語の媒介影響量の上位 10 語を表 3 に示し，比較のために出現頻度の上位 10 語を表 4 に示す．

表 3 語の媒介影響量の上位 10 語

順位	語	媒介影響量
1	店	0.172
2	色	0.158
3	フリース	0.145
4	ネット	0.145
5	買う	0.127
6	実物	0.127
7	ピンク	0.124
8	パールオリーブ	0.099
9	チラシ	0.061
10	ユニクロ	0.056

表 4 出現頻度の上位 10 語

順位	単語	出現頻度
1	色	26
2	買う	16
3	店	9
4	いい	7
4	働く	7
4	ピンク	7
4	店舗	7
4	見る	7
9	人気	6
10	セール	5

4・1 節，4・2 節での議論からも分かるように，実験に用いたコメントチェーンの話題の中心は，

- ピンク，パールオリーブ色のフリースが人気である．
- チラシとネットと実物でフリースの色は異なるのか？
- フルジップとハーフジップのフリース，どちらの方がかわいいか？

であった．したがって，媒介影響量により得られた語は話題の中心であった語が得られていることがわかる．一方，高頻度により得られた語は「働く」や「セール」など，周延的な話題であった語が含まれており，また「いい」「見る」は一般的に使われる語であるので，具体的にどのようなトピックで使われたのか見えてこない．したがって，高頻度の語より媒介影響量の大きい語の方が，話題の隆盛を的確に捉えていると言えよう．

5. 電子掲示板の分析事例 2

4 章では 17 件のコメントからなるコメントチェーンについて実験を行い，提案手法の有効性を検証した．そこで本章では，幾つものコメントチェーンが絡み合う 1161 件全てのコメントについて提案手法を適用し，電子掲示板のコメント中から盛り上がる話題を提供するコメント，オピニオンリーダー，語の発見を試みる．

5・1 盛り上がる話題を提供するコメントの発見

媒介影響量が上位のコメントは表 5 のようになり，それぞれのコメントの内容は次のようになる．コメント #1 はこの掲示板がこれほど大きくなったきっかけとなるコメントである．コメント #604 は 4 章で用いたコメントチェーンの起点となるコメントであり，4・1 節で述べたようにフリースに関する話題を提供して盛り上がったコメントである．コメント #460 はシャツや靴下などの直接肌に触れる商品の試着に関する話題について，コメント #1050 はユニクロで 50 歳まで働いたときに自分の居場所があるのかという話題を提供しており，実際にユニクロで働いているアルバイトや社員を巻き込んで盛り上がるきっかけとなったコメントである．コメント #807 はユニクロの製品は安いけど品質はどうなのかという内容であり，このコメントを起点として 10 ものコメントが熱い議論を繰り広げていた．

なお，発話の大きいコメントの上位は表 6 のようになるのだが，#807，#724，#799 はフレーミング，#1152

表 5 コメントの媒介影響量の上位 5 つ

順位	コメント ID	媒介影響量
1	#1	34.93
2	#604	1.504
3	#460	1.083
4	#1050	1.072
5	#807	1.071

は期間限定商品に関する話題，#1055 は就職活動に関する話題であり，いずれのコメントも盛り上がる話題を提供していなかった。

表 6 発話の大きいコメントの上位 5 つ

順位	コメント ID	発話の大きさ
1	#807	1433 文字
2	#1152	1100 文字
3	#724	1044 文字
4	#799	977 文字
5	#1055	969 文字

また，コメント #1，#604，#460，#1050，#807 の被返信数はそれぞれ 357，4，2，5，3 であり，被返信数が極めて多い #1 を除くと，コメントの媒介影響量は被返信数に因らないことがわかる。

以上のことから，幾つもの話題が混ざった電子掲示板の全てのコメントからも，媒介影響量を用いれば盛り上がる話題を提供しているコメントを的確に取り出せていることがわかる。

5.2 オピニオンリーダーの発見

媒介影響量が上位の人を表 7 に示す。M011 は #1 のコメントの発言者，M004 は #460 のコメントの発言者であり，それぞれのコメントの内容が場を盛り上げていることは 5.1 節で既に示した。M002，M010 はそれぞれユニクロのスタッフ（M002 は男性，M010 は女性）であり，着こなしの助言，お勧めの商品，店舗での実際の売れ筋などの情報を提供しているのみならず，他の参加者からアドバイスを請われる存在になっている。M021 もユニクロ歴 4 年のユニクロファンであり，情報収集に余念のない人である。いずれの人たちも多くのコメントの文脈を強く支配しているオピニオンリーダーである。

表 7 参加者の媒介影響量の上位 5 名

順位	メンバー ID	媒介影響量
1	M011	36.09
2	M002	3.949
3	M004	3.340
4	M010	2.985
5	M021	2.841

一方，被返信数の上位の人および発言数の上位の人をそれぞれ表 8，表 9 に示す。M002，M011 は既に述べたとおりオピニオンリーダーである。また，M009 も積極的に情報提供しているユニクロのスタッフであり，他のコメントの文脈を支配することが多いオピニオンリーダーである。しかし，M001 は批判的な発言が多い投稿者，M002，M004 は盛り上がる話題を提供している投稿者，M003，M005 は他者のコメントにまめに返信するが自分から新たな話題をあまり提供していない投稿者であった。したがって，被返信数，発言数が多い人がいつも話題を提供しているとは限らないことがわかる。

表 8 被返信数の上位 5 人

順位	メンバー ID	被返信数
1	M011	363
2	M001	40
3	M002	35
4	M003	25
5	M009	16

表 9 発言数の上位 5 人

順位	メンバー ID	発言数
1	M001	49
2	M002	44
3	M003	26
4	M004	23
5	M005	19

5.3 盛り上がる話題を提供する語の発見

媒介影響量の上位の語を表 10，出現頻度の多い語を表 11 に示す。この 2 つの表を見比べると，まず媒介影響量の上位 5 位までの語は全て出現頻度も 10 位以内であることが分かる。これらの語はユニクロの代表的な商品（「フリース」）とその特徴（「色」の種類が多い）を表す語であり，この掲示板を通して中心的な話題を提供していた。また「いい」「買う」「好き」「安い」「着る」はこれらの語に関係している場合が非常に多く，この掲示板の全体の志向性を的確に表している語である。

一方，表 10 でのみ得られた「カタログ」「ジーンズ」「返品」の頻度順位はいずれも 50 位以下であり，常に盛り上がっている話題ではなかった。しかし，電子掲示板の内容を詳細に検討してみると「カタログ」は店が混んでいるのが苦手な人や店に行く時間が取れない人たちに加え，カタログでしか買えない限定色に注目している人たちによって盛り上がった話題であった。また「ジーンズ」は年中販売されている商品であるので話題性は低いのだが，ジーンズ好きな人たちは流行に左右されずにジーンズの話で盛り上がっていた。「返品」に関しては，ユニ

表 10 語の媒介影響量の上位 10 語

順位	単語	媒介影響量
1	ユニクロ	15.02
2	いい	4.689
3	買う	2.926
4	フリース	1.677
5	色	1.355
6	カタログ	0.971
7	ジーンズ	0.969
8	好き	0.938
9	安い	0.934
10	返品	0.728

表 11 語の出現頻度の上位 10 語

順位	単語	出現頻度
1	ユニクロ	1021
2	フリース	252
3	買う	247
4	いい	197
5	着る	157
6	店	137
7	商品	136
8	色	134
9	服	128
10	見る	127

クロビジネス原則 [ファーストリテイリング 00] にある「理由を問わず、お買い上げから 3ヶ月間は、いつでも返品・交換します。」の「理由を問わず」というところに参加者の興味が集まり議論が繰り広げられていた。しかし、出現頻度の多い「店」「商品」「服」「見る」のような語はよく用いられる一般的な語であるために、盛り上がる話題を表しているとは言えない。

6. 議 論

4章, 5章での分析事例から, 発言数や被返信数などのコメントチェーンの構造や発話の大きさにだけ注目しては, 盛り上がる話題を発見することは難しいことがわかった。一方, 媒介影響量の尺度によって捉えられる話題は他のコメントたちの文脈を強く支配しており, そのような話題を表す語は頻度が低くても確かに議論を支配していたことがわかった。低頻度意見の方にユニークで面白い発言が多いことは従来から指摘されており [博報堂 00], 頻度が低くても他のコメントに合計で強く伝播されていれば取り出すことができる提案モデルは, ユーザの興味や価値観, ニーズを発掘するために有効であろう。

なお, 電子掲示板のように参加者の匿名性が高い場合には, 意見の極化, 議論への対等な参加, フレーミング (誹

謗中傷などの敵意的な言語行動) が起こりやすい [Kiesler 84]。しかし, フレーミングの引き金となるコメントには他のコメントの内容を支配するほどの影響力がないために, 文脈への支配関係に注目する提案モデルではフレーミングの影響を受けにくいという知見も得られた。

従来研究にはコメントチェーンの構造を社会ネットワーク分析 [安田 97] と呼ばれる手法で分析して, 有力なコメントや投稿者を発見することを狙った研究もある [金子 96, 北山 97]。しかし, 一連のコメントチェーンであってもそこで盛り上がる話題は変わっていくことも多いので, コメントの内容を追っていかないと真に意味のあるつながりは見えてこない。「>」などの引用符を手がかりにして内容の引用関係を見つけることも提案されているが [村上 00], 引用するかしないかは各人の書き癖によるところが大きい。したがって, 盛り上がる話題を発見するためには, 引用関係だけでなくコメントの内容まで考慮することが重要だと著者らは考えている。コメントの返信関係に沿って単語レベルで話題が広まっている程度を媒介影響量として定式化している理由はここにある。

7. ま と め

本論文では他のコメントの内容を強く支配する話題を盛り上がる話題と見なし, ある人の発した語への興味は他の人に伝播していくプロセスに着目して, テキストによるコミュニケーションにおける影響の普及を表すモデルを提案した。そして実際に電子掲示板に適用し, 盛り上がる話題を提供しているコメント, オピニオンリーダー, 語を取り出せることを示した。

なお, コメントチェーンの途中から新しい話題を始めるような変則的な展開は, その引き金をひいたコメントの間接的な結果と見ることもできる。そのような文脈の転換点を見出すことは社会の将来を考える予兆発見的な意義をもつので, 今後は話題の転換点に焦点を当てて研究を進めていきたい。

謝 辞

実験結果の評価方法について大変有益なご意見をいただいた査読者の方に感謝いたします。

◇ 参 考 文 献 ◇

- [池田 97] 池田謙一: ネットワーキングコミュニティ, 東京大学出版, 1997.
- [石川 00] 石川直人, コミュニティ戦略研究会: インターネットコミュニティ戦略, SOFTBANK, 2000.
- [Katz 55] E. Katz, P.F. Lazarsfeld: Personal Influence, The Free Press, 1955.
- [金子 96] 金子郁容・VCOM 編集チーム: 「つながり」の大研究, NHK 出版, 1996.
- [Kiesler 84] S. Kiesler, J. Siegel, and T.W. Mcguire: Social Psychological Aspects of Computer-Mediated Communication, American Psychologist, 39, pp. 1123-1134, 1984.

- [北山 97] 北山聡：電線交響主義，NTT 出版，1997．
 [国領 00] 国領二郎：デジタルネットワーク上における顧客間インタラクションによる情報価値の形成，ナレッジ・マネジメント学会ナレッジ・マネジメント研究年報第 2 号，pp. 9-19, 2000.
 [茶筌] 茶筌，<http://chasen.aist-nara.ac.jp/>
 [博報堂 00] 博報堂インタラクティブカンパニー：インターネットマーケティング，日本能率マネジメントセンター，2000．
 [ファーストリテイリング 00] ファーストリテイリング株式会社：ファーストリテイリング会社案内，2000．
 [村上 00] 村上明子，長尾確：ディスカッションマイニング：構造化されたコミュニケーションによるトピックの検索と視覚化，言語処理学会第 6 回年次大会発表論文集，pp. 451-454, 2000．
 [安田 97] 安田雪：ネットワーク分析，新躍社，1999．
 [吉田 00] 吉田豊：プロ野球関連掲示板におけるフレーミング現象，第 3 回 CMCC 研究会，2000.
 [Rogers 62] E.M. Rogers: Diffusion of Innovations, The Free Press, 1962.

〔担当委員：角 康之〕

2001 年 8 月 29 日 受理

—— 著 者 紹 介 ——



松村 真宏(学生会員)

1998 年大阪大学基礎工学部システム工学科卒業．2000 年同大学院修士課程修了．現在，東京大学大学院工学系研究科博士課程在学中．2000 年より科学技術振興事業団リサーチスタッフ．最近は人間の意思決定のプロセスに興味がある．2001 年人工知能学会 MYCOM 優秀プレゼンテーション賞受賞．



大澤 幸生(正会員)

1990 年東京大学工学部卒業．1995 年同大学院博士課程修了．博士(工学)．大阪大学基礎工学部助手を経て 1999 年より筑波大学社会学系助教授，現在に至る．2000 年より科学技術振興事業団研究者を兼任．予兆発見研究に従事．情報処理学会，AAAI，IEEE などの会員．人工知能学会では 1994 年，1999 年全国大会優秀論文賞，1998 年論文賞受賞．



石塚 満(正会員)

1971 年東京大学工学部電子卒業．1976 年同大学院博士課程修了．工学博士．同年 NTT 入社，横須賀研究所．1978 年東京大学生産技術研究所助教授．1992 年工学部電子情報工学科教授．2001 年より情報理工学系研究科電子情報学専攻．研究分野は人工知能，知識処理，マルチモーダル擬人化インターフェース/コンテンツ，WWW インテリジェンス．IEEE，AAAI，情報処理学会，人工知能学会，映像メディア学会，画像電子学会等の会員．